NBER WORKING PAPER SERIES

POLITICAL ACTIVISM AND THE PROVISION OF DYNAMIC INCENTIVES

Antoine Camous
Russell Cooper

Working Paper 26654
http://www.nber.org/papers/w26654

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

Political Activism and the Provision of Dynamic Incentives
Antoine Camous and Russell Cooper
NBER Working Paper No. 26654
January 2020
JEL No. D72,D74,E62,H31

ABSTRACT

This paper studies the determination of income taxes in a dynamic setting with human capital accumulation. The goal is to understand the factors that support an outcome without complete redistribution, given a majority of relatively poor agents. In the analysis, the internal dynamics of income are not sufficient to prevent complete redistribution under majority rule without commitment. However, a political influence game across the population limits the support for expropriatory taxation and preserves incentives. In some cases, the outcome of the game corresponds with the optimal allocation under commitment.

Antoine Camous
University of Mannheim
Department of Economics
L7 3-5 68131
Mannheim, Germany
camous@uni-mannheim.de

Russell Cooper
Department of Economics
European University Institute
Villa La Fonte
Via delle Fontanelle 18
I-50014 San Domenico di Fiesole (FI)
ITALY
and NBER
russellcoop@gmail.com

Political Activism and the Provision of Dynamic Incentives: Preserving the Pie in the Battle for Redistribution*

Antoine Camous[†], Russell Cooper[‡]

January 7, 2020

Abstract

This paper studies the determination of income taxes in a dynamic setting with human capital accumulation. The goal is to understand the factors that support an outcome without complete redistribution, given a majority of relatively poor agents. In the analysis, the internal dynamics of income are not sufficient to prevent complete redistribution under majority rule without commitment. However, a political influence game across the population limits the support for expropriatory taxation and preserves incentives. In some cases, the outcome of the game corresponds with the optimal allocation under commitment.

Keywords: *activism, electoral competition, commitment, redistribution, human capital.*

JEL classification: D72, D74, E62, H31.

1 Introduction

Income distributions are asymmetric, with a majority of agents earning below average income. Electoral competition under majority voting should implement tax and transfer schemes that are largely redistributive. In practice, labor income taxes are progressive but far from confiscatory. Accordingly, this paper asks: what factors constrain the relatively poor from expropriating the income and wealth of the relatively rich?

In a seminal contribution, Benabou and Ok (2001) provides two key conditions, based upon the *promise of upward mobility* (POUM), that alter the incentives of the relatively poor to tax the rich. First, a change in the tax structure must be permanent, or at least difficult to undo. Second, there must be sufficient mobility in the income distribution over time: the poor today recognize that they may be the rich of tomorrow. Hence, they are not in favor of high tax rates on the rich today since these same rates are likely to apply to them in the future.[1] If, to

*Thanks to Andrea Mattozzi, Guido Tabellini and seminar participants at University of Mannheim and European University Institute.

[†]Department of Economics, University of Mannheim, camous@uni-mannheim.de

[‡]Department of Economics, European University Institute, russellcoop@gmail.com

[1]A third necessary condition for POUM to hold is that individuals are not too risk averse so that poor agents find it attractive to be rich in the future.

the contrary, there is either insufficient mobility or sufficient flexibility in the tax system, then the relatively poor will favor the immediate and complete taxation of the rich.

The analysis of Benabou and Ok (2001) imposes an exogenous dynamic on the income distribution and thus misses adverse incentive effects stemming from redistribution. In our model, there is an explicit human capital decision that is forward looking and thus dependent on future taxation. This choice may be interpreted as formal education or more broadly as the accumulation of experience and the generation of ideas. The interaction of individual education decisions with the underlying distribution of ability implies that the dynamics of the income distribution do not satisfy the POUM conditions.[2]

Along with the consideration of incentives, timing matters: are taxes set prior to the human capital decision? If we maintain commitment to *ex ante* tax choices, as in Benabou and Ok (2001), then simple majority voting implements a social optimal level of taxes, not full redistribution. However, absent commitment, so that taxes can be set after the education choice, majority voting implements full redistribution, which in turns eliminates private incentives to invest in human capital.

Thus in an environment with endogenous mobility, no commitment and majority voting over taxes, the POUM hypothesis does not hold: the relatively poor will vote for complete redistribution. This first result motivates our analysis of alternative political institutions that might restrain full redistribution.

Our analysis deviates from standard electoral competition in two ways: probabilistic voting and activism. As in Persson and Tabellini (2002), probabilistic voting introduces stochastic elements into individual political preferences, interpreted as evaluations of politicians beyond their economic platforms. This framework allows other factors to influence voters' perceptions of a candidate, including the persuasive efforts of activism.

Activism embodies the idea that individuals, acting through groups, take joint actions to influence political outcomes. In contrast to elections that take place at discrete points in time, we view activism as a process that continuously influences political opinions, especially prior to the education decision.[3] This may take a variety of forms, ranging from direct political lobbying to public opinion persuasion through media and online campaigns.[4] Overall, activism by large income groups "distorts" the political ideal of one person, one vote.

We embed this political protocol in a dynamic model to study the interaction between human capital accumulation and these political institutions. Through the persuasion of voters, activism impacts the likelihood of redistribution, and in turn influences education choices and candidates' platforms: our results point to the social gains from a political system with activism under a lack of commitment.[5]

Indeed, when income groups organize as activists to influence the political preferences of voters, then the equilibrium outcome no longer coincides with full redistribution. As the analysis makes clear, there are conflicts across influence groups: rich income activists contribute for the low tax candidate while poor income agents contribute for the high tax one. But all influence groups internalize the benefits of preserving dynamic incentives. This provides a basis for relatively rich households to contribute more in favor of low taxes and, at the same time,

[2]In Benabou and Ok (2001), an essential condition for POUM to hold is that the income transition function displays negative skewness in the future: for the relatively poor median voter today to vote against redistribution tomorrow, it must be that she becomes richer than average tomorrow. Even with uncertainty, this is an unattractive description of income distributions. In our environment, the evolution of the income distribution is endogenous, but positive skewness is preserved.

[3]This timing is an important element of the analysis and is explained in detail below.

[4]See Becker (1983) for an early exposition of these ideas.

[5]To be clear, without human capital accumulation, our model is static and activism has no social value: full redistribution is the equilibrium outcome.

motivates the relatively poor to contribute less in favor of high taxes.[6] Overall, activism induces the relatively poor to vote for lower tax rates.

Candidates, in deciding on their policy platforms, internalize how activism will tilt voters preferences toward (or away from) their proposed redistribution rate. In equilibrium, activism eliminates the incentive for candidates to form a coalition around full redistribution. When the impact of activism is sufficiently strong on voter's preferences, the disciplining effect on candidates is powerful enough that the *ex ante* social optimum allocation with commitment is supported in equilibrium. As explained in detail, this outcome arises even though along the equilibrium path, the level of activism is zero by all groups. Thus, it is the credible threat of activism that matters.

Importantly, our environment does not impose any asymmetry across income groups in terms of participation or activism technology.[7] The outcome of less than full redistribution reflects the fact that all groups benefit from the dynamic incentives created by lower taxes. Activism provides a mechanism for the expression of these preferences.

Section 2 presents the economic environment. Section 3 derives a policy benchmark under commitment and discusses limits to political decentralization. Section 4 then studies how political activism shapes economic outcomes under electoral competition. All derivations and proofs are detailed in an Appendix.

2 Environment

Consider a two period $t = 1, 2$ economy populated by a continuum of agents. Agents at $t = 1$ differ in ability $\theta \sim \log \mathcal{N}(m, \sigma^2)$, with cumulative distribution function noted $F(\theta)$. They decide on education which influences next period's income.[8] At $t = 2$, agents are subject to productive idiosyncratic shocks. Taxes and transfers are applied to period 2 gross income, then agents consume net income. The economic channels of fiscal policy are multiple: it *redistributes* income across the population, provides *insurance* against idiosyncratic shocks and *distorts dynamic choices.*

2.1 Individual dynamic choice

An agent with ability θ at $t = 1$ invests in education e to maximize lifetime utility:

$$\max_{e} \ \log(c) + \beta E_{z,\tau}\big(\log(c')\big), \tag{1}$$

subject to:

$$c = \theta - e \qquad \text{budget constraint } t = 1 \tag{2}$$

$$\theta' = z\theta^{\alpha}e^{\delta} \qquad \text{gross income } t = 2 \tag{3}$$

$$c' = \theta'^{1-\tau}\bar{\theta}'^{\tau} \qquad \text{net income } t = 2 \tag{4}$$

[6] Importantly, the poor need to participate in the activism game to mitigate the low tax preferences of the rich.

[7] Prominent studies assume exogenous asymmetries to avoid outcomes with full redistribution: Benabou (2000) assumes exogenous differences in voting participation or Persson and Tabellini (2002) assume exogenous differential mass of swing voters across the population. Section 4.4 studies how asymmetries in activism technology influence our results.

[8] Throughout wages are set to unity so that human capital and income are the same.

Equation (2) is the budget constraint in period 1: gross income, equated to initial ability θ, is either consumed or spent on education. Equation (3) is the dynamic evolution of human capital and hence income. Both current ability θ and education choice e determine future income θ' up to an idiosyncratic shock $z \sim \log\mathcal{N}\big(-\frac{w^2}{2}, w^2\big)$. $\alpha > 0$ and $\delta \geq 0$ measure respectively depreciation of human capital and return to education.[9]

Fiscal policy in period 2 is summarized by (4). The isoelastic tax function is common in the literature on progressive labor taxes: the higher the redistribution rate $\tau \in [0, 1]$, the lower the dispersion of net income.[10] $\bar{\theta}'$ is a break-even income level which sorts the population in net beneficiaries and contributors. The expectation operator in (1) reflects uncertainty over the magnitude of taxes and transfers τ. The institutional structure that determines the period 2 tax rate will be a key element in the analysis.

The optimal education choice satisfies:

$$e(\theta, \bar{\tau}) = \epsilon(\bar{\tau})\theta, \tag{5}$$

where $\bar{\tau}$ is the expected average redistribution rate, and $\epsilon(\bar{\tau}) \equiv \frac{\beta\delta(1-\bar{\tau})}{1+\beta\delta(1-\bar{\tau})}$ is an education rate common to all agents. In the limit case of extreme redistribution, i.e. $\tau = \bar{\tau} = 1$, there is no private return to education, so that in period 2 agents have zero income and thus zero consumption.

Evolution of the income distribution. Gross income at $t = 2$ is log-normally distributed, with mean m' and standard deviation σ' given by:

$$m' = (\alpha + \delta)m + \delta \log\big(\epsilon(\bar{\tau})\big) - \frac{w^2}{2}, \tag{6}$$

$$\sigma'^2 = (\alpha + \delta)^2\sigma^2 + w^2. \tag{7}$$

Fiscal intervention. Through taxes and transfers, determined by τ, the fiscal intervention is purely redistributive. The critical income level $\bar{\theta}'$ sorts agents in net contributors $\theta' > \bar{\theta}'$ and net beneficiaries $\theta' \leq \bar{\theta}'$. It reads:

$$\log(\bar{\theta}') = m' + \frac{\sigma'^2}{2}(2 - \tau). \tag{8}$$

Note the dual influence of redistribution on $\bar{\theta}'$. First, average log-income m' is a function of education and expected tax rate $\bar{\tau}$, as explicit in (6). Second, $\bar{\theta}'$ is directly decreasing in τ: the share of the population that pays more taxes than it receives transfers is increasing in the magnitude of the redistributive program.

2.2 Value functions and bliss policies.

To highlight the influence of redistribution on individual preferences and welfare, we contrast individual bliss policies at two points in time: before agents form an education choice at $t = 1$, and after, at $t = 2$, when idiosyncratic uncertainty remains.[11]

[9] By design, individual education is not directly influenced by the human capital decisions of others. This allows to isolate the interaction of agents through fiscal policy.

[10] See for instance Benabou (2000), Benabou (2002) or Heathcote, Storesletten, and Violante (2017).

[11] This timing allows the exogenous component of mobility to shape individual preferences. It is maintained throughout the analysis.

At $t=1$, before agents choose education. Let $V_1(\theta,\tau)$ be the value function of an agent of type θ evaluating a rate of fiscal redistribution τ:

$$V_1(\theta,\tau) = \log\big(\theta - \epsilon(\tau)\theta\big) + \beta\Big((1-\tau)\big(\underbrace{\alpha\log(\theta) + \delta\log\big(\epsilon(\tau)\theta\big) - \frac{w^2}{2}}_{=E_z\big(\log(\theta')|\theta\big)}\big) + \tau\log(\bar{\theta}')\Big). \tag{9}$$

Here, there are two ways in which τ influences $V_1(\theta,\tau)$. First, the individual education decision $e(\theta,\tau)=\epsilon(\tau)\theta$ is sensitive to τ, as explicit in (5). Second, the government break-even income level $\bar{\theta}'$, given by (8), responds to the tax rate as well. While individual education choice (5) does not internalize the fiscal externality, individual evaluation of policy alternatives does.

The favorite redistribution rate $\tau^*(\theta)$ of a type θ agent is the tax rate $\tau\in[0,1]$ that solves $\frac{dV_1(\cdot)}{d\tau}=0$. From the first-order condition, it is implicitly given by:

$$\beta(\alpha+\delta)\big(m-\log(\theta)\big) + \beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \beta\delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = 0. \tag{10}$$

The first two terms in this expression capture the relative support for *redistribution*. At least all agents with income below (log) median level m benefits (on average) from redistributive policies. Also, an increase in individual risk w^2 generates higher desire for *insurance* via redistribution across the population. The third term is the elasticity of the education rate $\epsilon(\tau)$ to redistribution τ. It is negative and captures the willingness to preserve individual *dynamic incentives* against distortionary redistribution. It is straightforward to show that $\tau^*(\theta)<1$ is unique and decreasing in θ: higher income agents prefer lower rates of redistribution.[12]

At $t=2$, after individual choices. Let $V_2(\theta,\tau|\epsilon)$ be the value function of a type θ agent after $t=1$ consumption and education choices, before a realization of idiosyncratic risk z.[13]

$$V_2(\theta,\tau|\epsilon) = \beta\Big((1-\tau)\big(\underbrace{\alpha\log(\theta) + \delta\log(\epsilon\theta) - \frac{w^2}{2}}_{=E_z\big(\log(\theta')|\theta\big)}\big) + \tau\log(\bar{\theta}')\Big). \tag{11}$$

At that stage, education levels are no longer sensitive to tax policies τ: the education rate ϵ is a sufficient statistic to describe individual education $e=\epsilon\theta$ and the aggregate tax base. But, the break even income level $\bar{\theta}'$ is sensitive to τ insofar that it sorts agents between contributors and beneficiaries of the tax program, given average log income m'.

The favorite redistribution rate $\tau^d(\theta)$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau}=0$:

$$\beta(\alpha+\delta)\big(m-\log(\theta)\big) + \beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) = 0, \tag{12}$$

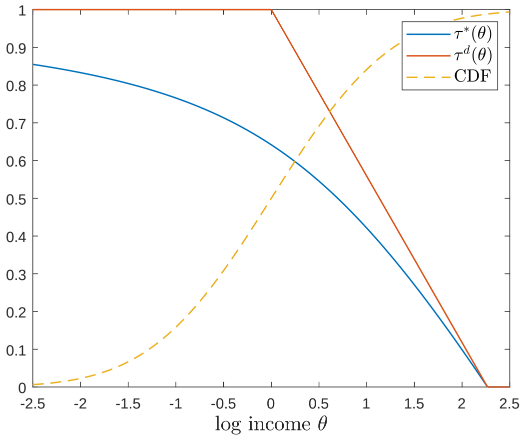
[12]This sharp characterization of individual preferences is provided by our assumptions on preferences and distributions. But the generic trade-offs embedded in the interactions between dynamic choices and redistributive policies would generalize to more general preference specifications.

[13]Given that the ranking of income before the realization of idiosyncratic uncertainty is maintained over time, we continue to order agents by initial income θ and CDF $F(\theta)$.

or the corner solution $\tau^d(\theta) = 1$ for lower income agents.[14] Further, as education is realized, there is no term capturing *dynamic incentives* effect of redistribution, in contrast to (10). This expression reflects otherwise the relative preferences for redistribution, decreasing with income, and desire of insurance against idiosyncratic shocks.

Figure 1 represents bliss policies before and after education, respectively $\tau^*(\theta)$ and $\tau^d(\theta)$. Favorite redistribution rates are decreasing in income θ. Agents internalizing the effect of taxes on incentives favor lower rates than they would after education is made: $\tau^*(\theta) \leq \tau^d(\theta)$ for all θ. Finally, after education, a majority supports complete redistribution.[15]

Figure 1: Individual Bliss policies



This figure represents bliss policies before and after education, as a function of (log) income. Favorite rates are decreasing in income and agents support higher levels of redistribution once education is made. The yellow dashed line shows the underlying distribution of log-income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1.]

3 Outcomes under Commitment: τ^*

Policy maker with commitment. Before formally defining political protocols, we characterize a key normative benchmark. Assume the policy choice is made by a benevolent policy maker at $t = 1$ with commitment. That is, taxes are chosen prior to the education choice and *ex post* are not subject to change.[16]

Formally, a benevolent policy maker with commitment chooses a tax rate τ^* to maximize expected utility over the population:

$$\max_{\tau} \int_{\theta} V_1(\theta, \tau) dF(\theta). \tag{13}$$

[14]Note that bliss policies $\tau^d(\theta)$ are only a function of income, not education ϵ: the size of the tax base is irrelevant and only redistributive conflicts determine bliss policies.
[15]The presence of positive skewness, as in the data of most if not all countries, is maintained throughout our analysis.
[16]In contrast taxes set without commitment are chosen in period 2, after the education decision.

Using (10), $\tau^* \in (0,1)$ is the solution to:

$$\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = 0. \tag{14}$$

The terms in this expression highlight how $\tau^* \neq 1$ balances (average) preferences for *redistribution / insurance* and *incentives*. The first term captures the insurance over ability θ and human capital risk z. The optimal rate of redistribution is increasing in income inequality σ^2 and idiosyncratic risk w^2.[17]

The second term captures the negative effect of higher taxes on human capital accumulation. The magnitude of this effect is parameterized by the return to education δ. But the overall effect of δ on τ^* is ambiguous, since it contributes both to higher return to education but also to income dispersion.

Political decentralization. This allocation can be decentralized through simple electoral competition with majority voting.[18] Interestingly, this does not depend on the timing of the vote relative to the realization of ability, θ. But it is critical that the vote is taken prior to education decisions.

First, suppose majority voting were to take place before the realization of ability. In this case, all agents are identical and their preferred tax policy would coincide with τ^* defined in (14). There is no conflict behind the veil of ignorance. Second, suppose the vote on the tax rate takes place after the realization of ability. Now there are well defined rich and poor agents. But key to the outcome is whether the vote takes place before or after education choices.

Proposition 1. *The allocation under τ^* can be decentralized under electoral competition with majority voting if voting takes place* ***before*** *the education choice. If, instead, voting takes place* ***after*** *education, then full redistribution is the outcome of electoral competition.*

In the first case majority voting does yield the efficient outcome τ^*, without full redistribution. The channel though differs from that identified in the POUM argument. In our environment with human capital accumulation, the median voter, in determining a preferred tax, internalizes the effects of taxes on the education choice of all others, and accordingly the size of the tax base m' for fiscal interventions, as explicit in (9). The critical element for this result is the timing: taxes are set prior to education choices.

In contrast, if majority voting were to happen after education, individual agents would no longer internalize the effect of policies on education, as discussed in Section (2.2). This yields an outcome of full redistribution, as it coincides with the bliss policy of a majority of the population, see Figure 1. In these circumstances, the anticipation of full redistribution eliminates private incentives to invest in human capital.

4 Probabilistic voting and activism

A key result from the previous section is that majority voting can support the efficient tax rate τ^* **iif** the vote is taken **prior** to the education choice. To the extent that education decisions are made before the determination of

[17]If redistribution over initial income θ were allowed, this would decrease initial income inequality and decrease the optimal rate of redistribution. But the commitment tension at the heart of the mechanism would be maintained.

[18]Simple electoral competition follows standard exposition in Persson and Tabellini (2002): two office seeking candidates propose competing policy platforms. An equilibrium policy survives pairwise evaluation to all possible policy alternative.

taxes, the challenge is to identify conditions such that incentives for human capital accumulation are preserved.

Our focus is on the political system. In our setting, candidates propose tax rates to maximize vote share. We augment the environment with a political protocol that combines pre-election politics and probabilistic voting. Probabilistic voting follows Persson and Tabellini (2002): agents evaluate candidates based upon their preferred tax rates and another dimension that reflects political preferences. Political activism has an impact on the voting outcome through this second dimension.

At the activism stage, agents with similar income levels organize as influence groups and decide cooperatively on non-pecuniary contributions to influence voters political preferences. Influence groups are motivated by the economic benefits of a policy platform. Activism takes place prior to the education choice and after the candidates have chosen their tax platforms. This reflects the idea that activism is an ongoing process in contrast to elections that take place at discrete points in time. Of course, candidates appreciate that their choice of policy platforms can elicit a response through the level and direction of political activism.

In equilibrium, activism matters and can push the outcome away from complete redistribution despite lack of commitment. The result rests upon two elements of the model. First, groups are large in that their actions can influence the voting outcome. Second activism disciplines candidates who would otherwise choose full redistributive tax policies. With this reduction in the probability of high taxes as the outcome of majority voting, agents retain an incentive to invest in human capital.

As the influence of activism on voters' preferences grow, the outcome of the game converges to the *ex ante* efficient outcome τ^*. Interestingly, there is no activism along the equilibrium path. The credible prospect of activism is enough to discipline candidates.

4.1 Timing of the game

The timing highlights both that human capital is determined prior to the vote on taxes and how activism influences the election outcome. Formally, the sequence of events is:

i. Choice of platforms: two office seeking candidates from competing parties L and H propose redistributive platforms $\tau_l \leq \tau_h$.

ii. Pre-election politics: activism.

iii. Individual choice at $t = 1$: agents, given their type θ, chooses consumption and education e.

iv. Political preferences: individuals are subject to idiosyncratic and aggregate political preference shocks.

v. Vote: given policy platforms and political preferences, agents participate in a majority election and the winning candidate takes office.

vi. Realizations of individual income shock z, tax and transfer and $t = 2$ consumption.

This timing calls for some comments. First, individual income uncertainty realizes after the vote, to give a chance, as in POUM, for *insurance* and *upward mobility* to influence the vote against the most redistributive policy platforms. Second, the vote takes place after the education choice, precisely to investigate whether highly redistributive platforms would emerge without commitment. Finally, activism takes place at a pre-election stage, allowing it to shape individual political preferences and in turn candidates policy platforms.

In order to isolate the effects of activism from the other elements of the political protocol, the first step is to characterize the political equilibrium of this game without activism. As we demonstrate, the outcome coincides with *ex post* simple majority rule: complete redistribution. Thus the results of Proposition 1 extend to a setting with probabilistic voting without activism.

4.2 No Activism

Absent pre-election politics, policy platforms (τ_l, τ_h) are decided by office seeking candidates anticipating the outcome of the probabilistic vote. The following exposition derives from the sequential nature of the game.

Voting outcome. Given policy platforms (τ_l, τ_h) and education rate ϵ, individuals trade off political and economic preferences and cast their vote sincerely. They evaluate policy platforms $\tau_l \leq \tau_h$ according to the value function $V_2(\theta, \tau|\epsilon)$ along with the realizations of idiosyncratic χ and aggregate ψ political preference shocks for candidate L. A type θ agent with education $e = \epsilon\theta$ votes for party H if and only if:

$$V_2(\theta, \tau_h|\epsilon) > V_2(\theta, \tau_l|\epsilon) + \chi + \psi. \tag{15}$$

As in Persson and Tabellini (2002), these shocks are distributed as:

$$\chi \sim U\Big(-\frac{1}{2\phi}, \frac{1}{2\phi}\Big) \qquad\qquad \psi \sim U\Big(-\frac{1}{2\Psi}, \frac{1}{2\Psi}\Big). \tag{16}$$

They differ only because the average of the idiosyncratic shock χ across the population is zero, while ψ is common across agents.[19]

Given a realization of aggregate preference ψ, let $\chi(\theta, \psi)$ be the *swing voter* for agents with income θ: type θ agents vote for party H if and only if $\chi \leq \chi(\theta, \psi)$. From (15),

$$\chi(\theta, \psi) = V_2(\theta, \tau_h|\epsilon) - V_2(\theta, \tau_l|\epsilon) - \psi = \Delta V_2(\theta) - \psi, \tag{17}$$

where $\Delta V_2(\theta)$ is the economic gain (or loss) to agents with initial income θ of τ_h over τ_l:

$$\Delta V_2(\theta) = \beta(\tau_h - \tau_l)\Big[(\alpha + \delta)\big(m - \log(\theta)\big) + \frac{\sigma'^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{18}$$

This expression makes clear that $\chi(\theta, \psi)$ and $\Delta V_2(\theta)$ do not depend on the actual education rate ϵ, as individual preferences are only driven by distributional conflicts at this stage. The vote share for party H within group θ and

[19] Though these shocks determine political preferences for candidate L, they impact the voting outcome symmetrically. As shown in Appendix C, the results not sensitive to the mean zero assumption, or to the strict majority requirement introduced below.

across the population are:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi}}^{\chi(\theta,\psi)} \phi dj = \phi\Big(\chi(\theta,\psi) + \frac{1}{2\phi}\Big) \qquad\qquad \pi_h(\psi) = \int_\theta \pi_{\theta,h}(\psi) dF(\theta). \tag{19}$$

In a majority system, the probability p_h that the candidate from party H wins the election is:

$$p_h = P\big(\pi_h(\psi) \geq 1/2\big). \tag{20}$$

Combining previous expressions:[20]

$$p_h = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) \geq \frac{1}{2}. \tag{21}$$

This expression highlights the tendency of the population to lean toward the most redistributive platform τ_h. When elections take place after the education choice, then a pure redistributive conflict drives economic preferences of agents. As illustrated in Figure 1, the positive skewness of the income distribution provides a majority mass of the population benefiting from high redistribution rates.

Choice of platforms. A candidate from party H seeking to maximize its probability of winning chooses to campaign on a redistributive program τ_h that solves:

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h), \tag{22}$$

where $p_h(\tau_l, \tau_h)$ is given by (21). The first order condition is:

$$\frac{dp_h(\cdot)}{d\tau_h} = \psi\beta\sigma'^2(1 - \tau_h) = 0. \tag{23}$$

$\tau_h = 1$ is a dominant strategy. The candidate from party L maximizes $p_l(\tau_l, \tau_h) = 1 - p_h(\tau_l, \tau_h)$, and again $\tau_l = 1$ is a dominant strategy.

Proposition 2. *The outcome of the game without pre-election politics is full redistribution $\tau^p = 1$.*

Intuitively, the election takes place after education choices have been formed. Hence despite redistributive conflicts, the median income agents with average political preferences support unconditionally a platform of full redistribution. In this case, dynamic incentives cannot be preserved and the economy collapses.

This corresponds to the outcome under *ex post* simple majority voting characterized in Proposition 1. Clearly, as in the static analysis of Persson and Tabellini (2002), probabilistic voting does not alter this outcome[21] But, once we introduce activism, the role of politics and probabilistic voting will be enhanced.

[20] Naturally $0 \leq p_h \leq 1$. We abstract from assumptions on parameters and restrictions on choices that are not relevant to characterize equilibrium outcomes.

[21] If the variance of idiosyncratic preference shocks is negatively correlated with income θ, then the outcome of electoral competition does not coincide with full redistribution: rich voters with more homogeneous political preferences over candidates have more swing voters, hence are more attractive targets to politicians. This is the essence of the analysis in Persson and Tabellini (2002), and is discussed in the context of our model in Section 4.4. Our analysis move beyond this exogenous source of asymmetry.

4.3 Electoral Competition with Activism

We now consider the equilibrium of the game when agents can actively support candidates and platforms.[22] By activism, we mean non-pecuniary contributions aimed to shape political preferences of voters. Which income groups contribute to the campaign of the high tax candidate? How does activism influence the probability of high taxes and equilibrium policy platforms? Eventually, does activism contribute to mitigate the high propensity of majority systems to lean toward redistributive policies?

We analyze activism intensity chosen by income groups and its impact on the voting outcome. The analysis takes as given group membership and focuses on the intensive margin of group specific contributions.[23] Groups are large and thus internalize the effects of their actions on the voting outcome.

Activism impact on the voting outcome. After the announcement of policy platforms (τ_l, τ_h), at the pre-election stage, each income group θ decides on (non-pecuniary) activism contributions $C_\theta^h \geq 0$ and $C_\theta^l \geq 0$ to promote candidates and their economic platforms. Contributions influence political preferences of voters.

At the time of the vote, these choices are all given. The vote reflects the economic valuations of policy platforms along with political preferences. Note $C^i = \int_\theta C_\theta^i d\theta$ the aggregated influence of activism for each candidate $i \in \{l, h\}$. Given $\left(C^l, C^h\right)$, aggregate preference shock ψ, education rate ϵ and policy platforms (τ_l, τ_h), an agent with initial income θ and preference shock χ votes for party H if and only if:

$$V_2(\theta, \tau_h|\epsilon) > V_2(\theta, \tau_l|\epsilon) + \chi + \psi + \gamma(C^l - C^h), \tag{24}$$

where $\gamma > 0$ measures how activism influences political preferences. The probability p_h that party H wins the election and τ_h is implemented reads:

$$p_h = \frac{1}{2} + \Psi \int_\theta \Delta V_2(\theta) dF(\theta) + \Psi\gamma(C^h - C^l). \tag{25}$$

Superficially, the term $\gamma(C^l - C^h)$ seems only to shift the distribution of the aggregate political preference shock. This misses an important dimension of the analysis: the choice of platforms is made anticipating the levels of activism. In this way, the prospect of political influence impacts the electoral platforms of candidates.

Choice of contributions. Every income group is active. The objective of each group is to influence political preferences of voters, with the goal of promoting their economic interests. In our setting, activism takes the form of effort rather than donations of consumption good. The costs of these contributions thus appear as utility costs.

Given all other group contributions $\{C_{-\theta}^l, C_{-\theta}^h\}$ and competing platforms $\boldsymbol{\tau} = (\tau_l, \tau_h)$, income group θ decides on total contributions $\left(C_\theta^l, C_\theta^h\right)$:

$$\max_{C_\theta^l, C_\theta^h \geq 0} f(\theta) V_1(\theta, \boldsymbol{\tau}) - \frac{1}{2}\Big((C_\theta^l)^2 + (C_\theta^h)^2\Big). \tag{26}$$

[22] The equilibrium concept is now two stage Nash equilibrium: first candidates choice of platforms and then activism effort.

[23] For expositional convenience, it is easier to work with a continuum of agents where each ability type is a distinct group. This allows to highlight differential activism incentives across the income distribution. The same results would go through if agents would form two groups campaigning exclusively for one candidate.

Here $V_1(\theta, \boldsymbol{\tau})$ is the expected value of a type θ household prior to the choice of education and to the election outcome for given political platforms.[24] It is similar to (9) but captures uncertainty over the tax rate $\tau \in (\tau_l, \tau_h)$:

$$V_1(\theta, \boldsymbol{\tau}) = E_\tau \left\{ \log\big(\theta - \epsilon(\bar{\tau})\theta\big) + \beta\big((1-\tau)\alpha\log(\theta) + \delta\log(\epsilon(\bar{\tau})\theta) - \frac{w^2}{2} + \tau\log(\bar{\theta}')\big) \right\}, \tag{27}$$

Activism internalizes the effect of contributions on the education rate $\epsilon(\bar{\tau})$ and the outcome of the vote, captured by the probability p_h that party H defeats party L, see (25). Income break even level $\bar{\theta}'$ is a function of both $\bar{\tau}$ and τ, as explicit in (8). The group cost in (26) is assumed to be quadratic. One interpretation is there is a cost of organization associated with activism that depends on the contributions of all members of the group [25]

The first order condition for $C^i_\theta \geq 0$, $i \in (l, h)$, is

$$f(\theta)\frac{dV_1(\theta)}{dC^i_\theta} = C^i_\theta. \tag{28}$$

The sensitivity of group θ welfare to activism is:[26]

$$\frac{dV_1(\cdot)}{dC^i_\theta} = \pm\Psi\gamma\beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m_1 - \log(\theta)\big) + \frac{\sigma_2^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{29}$$

The first two terms reflects the relative position of group θ in the income distribution and the preferences for redistribution, as in (10).

The last term reflects the effects of group contributions on the probability of high taxes and, through the expected tax rate, on the common accumulation of human capital $\epsilon(\bar{\tau})$.[27] This term is central to understand the outcome of the activism game: regardless of an agent's position in the income distribution, this effect pushes in the direction of low taxes since all agents prefer a large tax base.

This last term is also the locus of strategic interactions, i.e. **conflicts** across the groups. Again, the decision to support a low or high tax candidate as well as the magnitude of group contributions depend on the probability p_h, which hinges on overall activism across the population, as explicit in (25). As long as $\tau_l < \tau_h$, all groups are active, and contribute only for a single candidate: low income agents support only the champion of high taxes, while high income agents only the candidate from party L. Importantly, the split of the population is endogenous. All groups with initial income $\theta < \hat{\theta}$ contribute exclusively for τ_h, where the cut-off income level $\hat{\theta}$ is given by:

$$\log(\hat{\theta}) = m + \frac{1}{\alpha+\delta}\Big[\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{30}$$

Figure 2 report activism per capita $c^i_\theta = C^i_\theta / f(\theta)$ and total group activism C^i_θ for some policy platforms $\tau_l < \tau_h$. Individual activism is increasing in the income difference $|\hat{\theta} - \theta|$. Aggregate group effort is not monotonic, because

[24] The program (26) omits uncertainty regarding future political taste shocks. This does not change the analysis beyond overburdening notations.

[25] This specification is taken from Persson and Tabellini (2002). If the individual cost to agent j with income θ of contributing $c^{i,j}_\theta$ is $\frac{c^{i,j}_\theta}{2} c^i_\theta f(\theta)$, where c^i_θ is the average group contribution, then total group costs is $\frac{{C^i_\theta}^2}{2} = \frac{(c^i_\theta f(\theta))^2}{2}$: individual cost is linear, but increases with group size and average contribution.

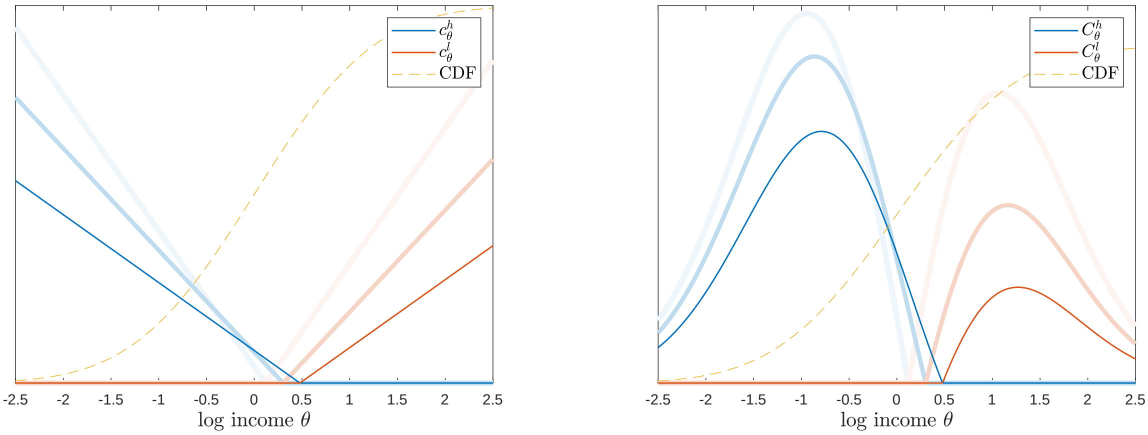
[26] In this expression, $\pm = \mathbb{1}_{i=h} - \mathbb{1}_{i=l}$

[27] Formally, $\bar{\tau} = p_h\tau_h + (1-p_h)\tau_l$.

of the relative size $f(\theta)$ of each income group. Finally, an increase in activism technology γ has two effects: it increases returns to activism but modifies the cut-off $\hat{\theta}$, i.e. it changes the composition of the population that promotes one candidate or the other. Aggregating group contributions (28), one gets:

$$C^h - C^l = \Psi\gamma\beta(\tau_h - \tau_l)\Big[\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{31}$$

Figure 2: Activism contribution



(a) Activism per capita (b) Total group activism

This figure represents activism per capita (left panel) and total group contributions (right panel) given two policy platforms $\tau_l = 1/4$ and $\tau_h = 3/4$. Lighter lines correspond to higher values of γ. Dashed yellow line reflects the distribution of (log) income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 1. *Given (τ_l, τ_h), there is a unique Nash equilibrium of the activism subgame. The induced probability that the high tax candidate wins the election is given by:*

$$p_h(\tau_l, \tau_h) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\Big[(1 + \Psi\gamma^2)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \Psi\gamma^2\delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{32}$$

The probability of high taxes takes into account the aggregation of group activism (31) and reflects the resolution of conflicts over redistribution. Importantly, this expression highlights the interactions of activism and human capital accumulation in reducing the probability of high tax rates. In the absence of a return to education, i.e. $\delta = 0$, $p_h \geq \frac{1}{2}$ as in (21). Only when $\delta > 0$ and $\gamma > 0$ does the last term in (32) reduce the probability of high taxes.

Choice of platforms. How does activism influence political competition and equilibrium incentives? At the initial stage of the game, each candidate decides on its economic policy platform anticipating the effect of activism on voters' political preferences and the outcome of the vote. Formally, candidate from party H sets τ_h given τ_l to maximize (32):

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h). \tag{33}$$

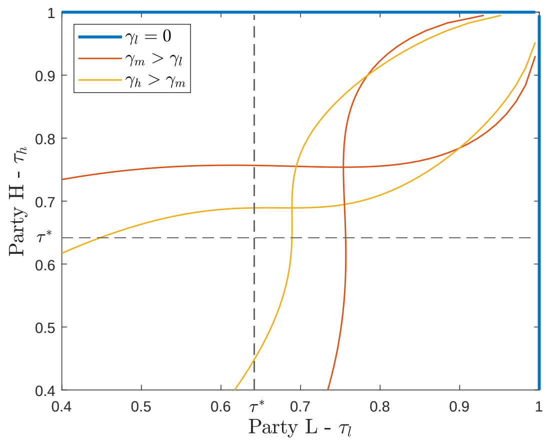
The first order condition leads to:

$$(1+\Psi\gamma^2)\sigma'^2(1-\tau_h) + \Psi\gamma^2\delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} + (\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}p_h(\tau_l,\tau_h) = 0, \tag{34}$$

where $\mathcal{E}(\bar{\tau}) = \bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} < 0$ is the elasticity of the education rate to the expected redistribution rate.

The platform choice internalizes the effect of the tax rate on the outcome of the activism subgame. Again, it highlights the interplay between activism and dynamic incentives. In the absence of dynamic choice, i.e. when $\delta = 0$, activism and associated conflicts are irrelevant since $\tau_h = 1$ is a dominant strategy, as in (23). In contrast, when $\delta > 0$, then activism induces strategic interactions across candidates. Figure (3) represents the best response functions of each candidate for different level of activism technology γ.

Figure 3: Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, without activism ($\gamma_l = 0$) and with varying degrees of activism ($\gamma_h > \gamma_m > 0$). [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 2. *There is a unique and symmetric Nash equilibrium $\tau^p = \tau_l = \tau_h$ that satisfies:*

$$\Big(1 + \frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1-\tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} = 0. \tag{35}$$

Activism does not imply that parties have an incentive to differentiate from each other in equilibrium, but as long as there are positive returns to education $\delta > 0$, the equilibrium no longer coincides with full redistribution.

Proposition 3. *The equilibrium rate of redistribution τ^p differs from $\tau = 1$ if and only if $\delta > 0$ and $\gamma > 0$. Further, it is decreasing in activism intensity γ, and in the limit case where $\gamma = +\infty$, the political game with activism implements the socially desirable level of redistribution τ^*.*

Along the equilibrium path of this game, there is no political activism since $\tau_l = \tau_h = \tau^p$. Still activist groups stand ready to influence voters political preferences if a candidate were to deviate from the equilibrium platform. Also, the equilibrium platform reflects the relative power of influence groups. If, for instance, the candidate from party L deviates from τ^p and runs on $\tau_l < \tau^p$, then relatively poor agents would engage into activism to convince voters away from this platform. And vice versa. This is the disciplining effect of activism on candidates.

The intensity of the activism technology γ influences the equilibrium rate of redistribution τ^p via the conflictual behavior across groups: for higher γ, the disciplining effect of activism is stronger. Again, the response of education to taxes, parameterized by δ is necessary for this channel.

Overall, activism is a complementary institution to majority voting to implement socially desirable policies in a dynamic environment with lack of commitment. It matters because agents, regardless of their position in the income distribution, appreciate the social benefit of human capital accumulation. So, as long as there is a response of education to activism's effect on expected taxes, $\delta > 0$, the bias towards redistribution under simple majority voting is, at least partially, redressed.

4.4 Extensions

This section studies two key extensions. What is the effect of asymmetry in activism technology on equilibrium outcomes? And, how do office-seeking candidates also interested to the policy outcome interact?

Asymmetric influence. The equilibrium outcome with activism characterized in Proposition 3 does not rely on exogenous asymmetry across groups to move away from full redistribution. In contrast, Persson and Tabellini (2002) consider asymmetry in the mass of swing voters across the population as a necessary condition to avoid full redistribution: when high income agents are more attractive targets to candidates, this reduces the likelihood of high taxes.[28]

In this section, we show that the introduction of asymmetry in activism technology might further reduce the likelihood of high taxes. Formally, we consider a situation where income groups differ in their technology to influence political preferences. The cost of contributing is decreasing in α_θ, with $\int_\theta \alpha_\theta d\theta = 1$. Given policy platforms (τ_l, τ_h) and other groups' contributions $\{C^l_{-\theta}, C^h_{-\theta}\}$, the activism choice of income group θ is the solution to:

$$\max_{C^l_\theta, C^h_\theta \geq 0} f(\theta) V_1(\theta, \boldsymbol{\tau}) - \frac{1}{2\alpha_\theta}\Big((C^l_\theta)^2 + (C^h_\theta)^2\Big). \tag{36}$$

Let $\rho = \text{cov}\big(\alpha_\theta, \log(\theta)\big)$ be the covariance between income level and influence technology. The first order condition for $C^i_\theta \geq 0$ then reads:

$$\alpha_\theta f(\theta) \frac{dV_1(\theta, \boldsymbol{\tau})}{dC^i_\theta} = C^i_\theta. \tag{37}$$

Group contribution is increasing in influence technology α_θ. As $E\big(\alpha_\theta \log(\theta)\big) = m_1 + \rho$, the probability $p_h(\tau_l, \tau_h)$

[28]Formally, in our set up assume $\chi_\theta \sim U\Big(-\frac{1}{2\phi_\theta}, \frac{1}{2\phi_\theta}\Big)$: if $\text{cov}(\phi_\theta, \log(\theta)) < 0$, then the outcome of the political game without activism yields $\tau^p < 1$. See Appendix C.1 for explicit derivations.

of party H win is then implicitly defined by:

$$p_h(\tau_l,\tau_h) = \frac{1}{2} + (\Psi\gamma)^2\beta(\tau_h-\tau_l)\Big[(1+\frac{1}{\Psi\gamma^2})\frac{\sigma'^2}{2}(2-\tau_h-\tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha+\delta)\rho\Big]. \tag{38}$$

A Nash equilibrium of the game across office seeking candidates $\tau^p = \tau_l = \tau_h$ is then the solution to:

$$\Big(1+\frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1-\tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha+\delta)\rho = 0. \tag{39}$$

In that context, $\frac{d\tau^p}{d\rho} < 0$: the larger the influence of high income groups, the lower the equilibrium rate of redistribution. Accordingly, a positive covariance between income level and activism technology can either compensate a low aggregate activism technology γ and bring the equilibrium level of redistribution toward the optimal level or can tilt the policy rate towards partisan level of redistributions that are too low.

Citizen-office-seeking candidates. So far, candidates were simply interested in being elected. We now allow candidates to obtain utility from office but also to be interested in the policy outcome. Specifically, they suffer a loss that depends on the distance between their proposed tax and their bliss policy.[29] This has two effects on the equilibrium outcome. First, the proposed taxes by the two candidates are no longer equal. Second, there is activism in equilibrium.

Formally, the candidate from party H seeks to maximize the probability of getting elected, but would also like his bliss policy τ_h^* to be implemented in equilibrium. Given τ_l, this candidate solves the following program:

$$\max_{\tau_h} \mu p_h + (1-\mu)E_\tau\Big[\frac{\big(\tau-\tau_h^*\big)^2}{2}\Big], \tag{40}$$

where p_h is given by (32). $0 < \mu < 1$ is a preference weight for being elected relative to the loss incurred as the actual policy rate deviates from the candidate's bliss policy. The candidate from party L solves a similar program with bliss policy τ_l^*.

Figure 4 illustrates candidates' best responses and the equilibrium outcome: $\tau_l^p < \tau_h^p$ since candidates offer policy platforms that are now influenced by their individual preferences $\tau_l^* < \tau_h^*$. As explained in Section 4.3 and explicit in (31), different policy platforms are associated in equilibrium with conflictual activism.

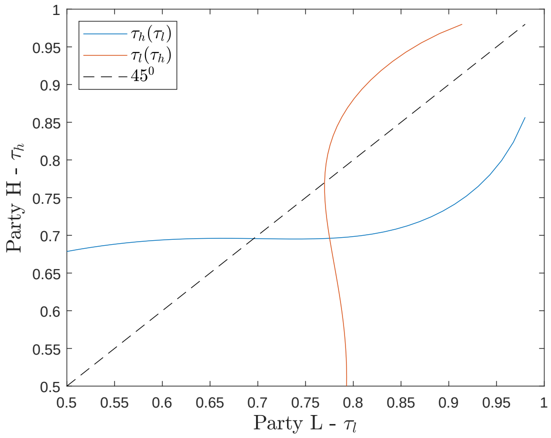
5 Conclusion

Empirically, labor taxes are progressive but do not expropriate all the earning of the rich to compensate the poor. Instead, redistribution is limited. This paper studies the economic and political factors that limit redistribution in a democracy.

In the environment, the arguments put forth by Benabou and Ok (2001) do not hold: absent commitment and despite the potential for upward mobility in the income distribution, the outcome under majority voting would be complete redistribution. In our model, this implies no capital accumulation and thus a massive reduction in the

[29]Discussions with Annika Bacher led to the development of this case.

Figure 4: Citizen-Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, with $\mu = 1/2$, $\tau_l^* = 0$, $\tau_h^* = 1$. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

"economic pie".

The analysis provides another mechanism: the power of persuasion. Coalitions of agents jointly decide on the level of activism which can influence the political preferences of agents and thus voting outcomes. Majority voting remains but the progress of political persuasion facilitates a redistribution of political power. Though there is no commitment, the outcome with activism is closer to the efficient allocation: redistribution is incomplete and incentives are retained for the accumulation of human capital.

References

Becker, G. S. (1983): "A theory of competition among pressure groups for political influence," The quarterly journal of economics, 98(3), 371–400.

Benabou, R. (2000): "Unequal Societies: Income Distribution and the Social Contract," American Economic Review, 90(1), 96–129.

——— (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," Econometrica, 70(2), 481–517.

Benabou, R., and E. A. Ok (2001): "Social Mobility and the Demand for Redistribution: The Poum Hypothesis," The Quarterly Journal of Economics, 116(2), 447–487.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework*," The Quarterly Journal of Economics, 132(4), 1693–1754.

Persson, T., and G. Tabellini (2002): Political Economics: Explaining Economic Policy, vol. 1. The MIT Press, 1 edn.

Mathematical Appendix

A Section 2 - Environment

Individual choice. Consider household optimization (1), substitute the constraints into the objective function:

$$\max_{e} \log(\theta - e) + \beta \int_{\tau} \Big\{(1-\tau)\big[\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}\big] + \tau \log(\bar{\theta}')\Big\} dG(\tau), \tag{41}$$

where $G(\cdot)$ captures uncertainty over rate τ. The first order condition reads:

$$-\frac{1}{\theta - e} + \frac{\beta\delta(1-\bar{\tau})}{e} = 0, \tag{42}$$

where $\bar{\tau} = E(\tau)$. Reorganize and get (5).

Evolution of the income distribution. Start from (3), take the log and use (5):

$$\log(\theta') = \log(z) + (\alpha + \delta)\log(\theta) + \delta \log\big(\epsilon(\bar{\tau})\big) \tag{43}$$

The mean and variance of this expression yield (6) and (7).

Break-even income level. Given τ, income net of taxes and transfers satisfies (4), take the integral and then the log:

$$\log \int_{\theta} c' dF(\theta) = \tau \log(\bar{\theta}') + \log \int_{\theta} {\theta'}^{1-\tau} dF(\theta). \tag{44}$$

If $X \sim \log\mathcal{N}(\mu, \sigma^2)$, then $E(X^n) = e^{n\mu + n^2\sigma^2/2}$. Since $E(c') = E(\theta')$, it gives:

$$m' + \frac{{\sigma'}^2}{2} = \tau \log(\bar{\theta}') + (1-\tau)m' + \frac{(1-\tau)^2}{2}{\sigma'}^2. \tag{45}$$

Reorganize and get (8).

Value functions and bliss policies. Lifetime utility to an agent with initial income θ at $t = 1$ reads:

$$\begin{aligned} V_1(\theta, \tau) &= v_1(\theta, e, \tau, \bar{\theta}') \\ &= \log(\theta - e) + \beta\big[(1-\tau)\underbrace{\big(\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}\big)}_{=E_z\left(\log(\theta')|\theta\right)} + \tau \log(\bar{\theta}')\big], \end{aligned} \tag{46}$$

where $e = \epsilon(\tau)\theta$ as in (5) and $\bar{\theta}'$ is given by (8). The first order condition w.r.t. τ:

$$\frac{dV_1(\cdot)}{d\tau} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e}}_{=0}\frac{de}{d\tau} + \frac{\partial v_1(\cdot)}{\partial \tau} + \frac{\partial v_1(\cdot)}{\partial \log \bar{\theta}'}\frac{d\log\bar{\theta}'}{d\tau} = 0. \tag{47}$$

The first term is 0 from the envelope condition, the other terms are:

$$\frac{\partial v_1(\cdot)}{\partial \tau} = \log(\bar{\theta}') - \big(\alpha\log(\theta) + \delta\log(e) - \frac{w^2}{2}\big), \tag{48}$$

$$\frac{d\log(\bar{\theta}')}{d\tau} = \frac{dm'}{d\tau} - \frac{\sigma'^2}{2} = \delta\frac{\epsilon'(\cdot)}{\epsilon(\cdot)} - \frac{\sigma'^2}{2}. \tag{49}$$

Reorganize and get (10):

$$\beta(\alpha+\delta)\big(m - \log(\theta)\big) + \beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \beta\delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = 0. \tag{50}$$

This expression implicitly defines $\tau^*(\theta)$. The first two terms form a decreasing linear function of τ, whose intercept is decreasing in θ. The third term, the elasticity of the saving rate to the redistribution rate τ, is decreasing in τ. Formally:

$$\frac{d}{d\tau}\Big(\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau}\Big) = -\Big[\frac{\tau}{1-\tau}\frac{\beta\delta}{\big[1+\beta\delta(1-\tau)\big]^2} + \frac{1}{(1-\tau)^2}\frac{1}{1+\beta\delta(1-\tau)}\Big] < 0, \tag{51}$$

which goes to $-\infty$ when τ goes to 1. Altogether there is a unique solution $\tau^*(\theta) < 1$ to (50), decreasing in θ.

The value function to an agent with initial income θ after $t = 1$ consumption and education choice reads:

$$\begin{aligned} V_2(\theta,\tau|\epsilon) &= v_2(\theta, e, \tau, \bar{\theta}'|\epsilon) \\ &= \beta\big[(1-\tau)\Big(\alpha\log(\theta) + \delta\log(e) - \frac{w^2}{2}\Big) + \tau\log(\bar{\theta}')\big] \end{aligned} \tag{52}$$

The difference with (46) is that the education choice is no longer sensitive to the redistribution rate τ and $e = \epsilon\theta$. The sensitivity of individual value functions to τ satisfies:

$$\begin{aligned} \frac{dV_2(\cdot)}{d\tau} &= \frac{\partial v_2(\cdot)}{\partial \tau} + \frac{\partial v_2(\cdot)}{\partial \log\bar{\theta}'}\frac{d\log\bar{\theta}'}{d\tau} \\ &= \beta(\alpha+\delta)\big(m - \log(\theta)\big) + \beta(\alpha+\delta)^2\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) \end{aligned} \tag{53}$$

Bliss policy $\tau^d(\theta) \in [0,1]$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau} = 0$ or $\tau^d(\theta) = 1$. Given the linear nature of (53), there is a unique bliss policy, ordered by initial income θ.

B Section 3 - Outcome under Commitment

Sensitivity of optimal redistribution rate. To derive comparative statics for τ^*, first get

$$\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = -\frac{\tau}{1-\tau}\frac{1}{1+\beta\delta(1-\tau)}. \tag{54}$$

Then rewrite (14) as

$$\big((\alpha+\delta)^2\sigma^2 + w^2\big)(1-\tau)^2 - \frac{\delta\tau}{1+\beta\delta(1-\tau)} = 0 \tag{55}$$

The total derivative of this expression:

$$\begin{aligned}(\alpha+\delta)^2(1-\tau)^2 d\sigma^2 + (1-\tau)^2 dw^2 + \Big[2(\alpha+\delta)(1-\tau)^2 - \tau\frac{1+2\beta\delta(1-\tau)}{\big(1+\beta\delta(1-\tau)\big)^2}\Big]d\delta \\ = \Big[2\sigma'^2(1-\tau) + \delta\frac{1+\beta\delta}{\big(1+\beta\delta(1-\tau)\big)^2}\Big]d\tau\end{aligned} \tag{56}$$

Get immediately $\frac{d\tau^*}{d\sigma^2} > 0$ and $\frac{d\tau^*}{dw^2} > 0$ while the sign of $\frac{d\tau^*}{d\delta}$ is ambiguous.

Proof Proposition 1. Individual value functions are single peaked and bliss policies ordered by income level. Pairwise evaluation of policy alternative leads to a convergence of electoral platforms toward the bliss policy of the median agent, with (log) income $m = \log(\theta^m)$. If the vote takes place before the education choice, then from (10), $\tau^*(\theta^m)$ coincides with τ^* defined by (14). If the vote takes place after education, then from (12) the outcome of the vote is full redistribution $\tau^d(\theta^m) = 1$.

C Section 4 - Probabilistic Voting and Activism

C.1 No Activism.

Comparison of platforms (18)**.** Start from (11):

$$V_2(\theta,\tau|\epsilon) = \beta\Big((1-\tau)\big(\alpha\log(\theta) + \delta\log(\epsilon\theta) - \frac{w^2}{2}\big) + \tau\log(\bar\theta')\Big), \tag{57}$$

with $\log(\bar\theta')$ given by (8). Rearrange and get:

$$V_2(\theta,\tau|\epsilon) = \beta\Big[(1-\tau)(\alpha+\delta)\log(\theta) + \delta\log(\epsilon) - \frac{w^2}{2} + \tau(\alpha+\delta)m + \frac{\sigma'^2}{2}(2-\tau)\tau\Big]. \tag{58}$$

The difference of this expression with τ_h and τ_l:

$$\Delta V_2(\theta) = \beta(\tau_h-\tau_l)(\alpha+\delta)\big(m-\log(\theta)\big) + \beta\frac{\sigma'^2}{2}\big[(2-\tau_h)\tau_h - (2-\tau_l)\tau_l\big]. \tag{59}$$

Verify $(2-\tau_h)\tau_h - (2-\tau_l)\tau_l = (\tau_h - \tau_l)(2-\tau_h-\tau_l)$ and get (18). Finally, note

$$\int_\theta \Delta V_2(\theta) dF(\theta) = \beta(\tau_h - \tau_l)\frac{\sigma'^2}{2}(2-\tau_h-\tau_l). \tag{60}$$

Probability p_h of high rate of redistribution τ_h. Consider the following distributions: $\chi^j_\theta \sim U\Big(-\frac{1}{2\phi_\theta}+m, \frac{1}{2\phi_\theta}+m\Big)$ and $\psi \sim U\Big(-\frac{1}{2\Psi}+M, \frac{1}{2\Psi}+M\Big)$. Note $\phi = E(\phi_\theta)$. Given ψ and competing platforms (τ_l, τ_h), agent j with income θ votes for party H if and only if $\chi^j_\theta \leq \chi(\theta,\psi) = \Delta V_2(\theta) - \psi$. Hence, the share of agents with income θ that vote for party H is:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi_\theta}+m}^{\chi(\theta,\psi)} \phi_\theta dj = \phi_\theta\Big(\chi(\theta,\psi) + \frac{1}{2\phi_\theta} - m\Big). \tag{61}$$

The share of votes across groups is then $\pi_h(\psi) = \int_\theta \pi_{\theta,h}(\psi)dF(\theta)$:

$$\pi_h(\psi) = \int_\theta \phi_\theta\Big(\chi_\theta + \frac{1}{2\phi_\theta} - m\Big)dF(\theta), \tag{62}$$

$$= \int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - \phi(\psi + m) + \frac{1}{2}. \tag{63}$$

The probability that party H wins the election is $p_h = P\big(\pi_h(\psi) \geq \lambda\big)$, where $\lambda \in [0,1]$ is a majority requirement for party H to win the election. The event $\pi_h(\psi) \geq \lambda$ is equivalent to the event

$$\psi \leq \bar{\psi} = \frac{1}{\phi}\int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - m + \frac{1}{\phi}\big(\frac{1}{2} - \lambda\big). \tag{64}$$

Get then p_h as

$$p_h = P\big(\psi \leq \bar{\psi}\big) = \frac{1}{2} + \frac{\Psi}{\phi}\int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - \Psi(m+M) + \frac{\Psi}{\phi}\Big(\frac{1}{2} - \lambda\Big). \tag{65}$$

Set $\phi_\theta = \phi$, $\lambda = \frac{1}{2}$, $m = M = 0$ and using (60) get (21).

Asymmetric political preferences. Persson and Tabellini (2002) study probabilistic voting allowing for a correlation between income and political preferences. To see how this mechanism works without activism, normalize average political preferences heterogeneity $\phi = E(\phi_\theta) = 1$ and assume that $\text{cov}\big(\phi_\theta, \log(\theta)\big) = a$: if $a > 0$ then high income agents are more responsive to economic factors. The probability $p_h(\tau_l, \tau_h)$ is then

$$p_h(\tau_h, \tau_l) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\Big(-(\alpha+\delta)a + \frac{\sigma'^2}{2}(2-\tau_h-\tau_l)\Big), \tag{66}$$

and a Nash equilibrium of economic platforms competition differs from full redistribution if and only if $a > 0$:

$$\tau_h = \tau_l = 1 - \frac{(\alpha+\delta)a}{\sigma'^2} < 1. \tag{67}$$

Overall, when high income agents are more sensitive to economic policy rather than political factors, they are attractive targets to candidates, which tilts equilibrium economic platform toward lower rate of redistribution.

C.2 Activism

Choice of contributions. To get (29), rewrite (27) as:

$$V_1(\theta, \boldsymbol{\tau}) = v_1\big(\theta, \boldsymbol{\tau}, e, p_h, \log(\bar{\theta}'_h), \log(\bar{\theta}'_l)\big) \tag{68}$$

where $e = \epsilon(\bar{\tau})\theta$. Further, $\frac{dV_1(\cdot)}{dC^i_\theta} = \frac{dV_1(\cdot)}{dp_h}\frac{dp_h}{dC^i_\theta}$. Consider the first term:

$$\frac{dV_1(\cdot)}{dp_h} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e}}_{=0}\frac{de}{dp_h} + \frac{\partial v_1(\cdot)}{\partial p_h} + \sum_{i=l,h}\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)}\frac{d\log(\bar{\theta}'_i)}{dp_h}. \tag{69}$$

Term by term:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = -\beta(\tau_h - \tau_l)E_z\big(\log(\theta')|\theta\big) + \beta\big(\tau_h\log(\bar{\theta}'_h) - \tau_l\log(\bar{\theta}'_l)\big), \tag{70}$$

and

$$\tau_h\log(\bar{\theta}'_h) - \tau_l\log(\bar{\theta}'_l) = (\tau_h - \tau_l)\Big[m' + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{71}$$

Using (6) and $E_z\big(\log(\theta')|\theta\big) = (\alpha+\delta)\log(\theta) + \delta\log\big(\epsilon(\bar{\tau})\big) - \frac{w^2}{2}$, rearrange and get:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = \beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{72}$$

Then, since $\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)} = \beta p_i \tau_i$ and $\frac{d\log(\bar{\theta}'_i)}{dp_h} = \delta\frac{d\bar{\tau}}{dp_h}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ for $i \in (l, h)$,

$$\sum_{i=l,h}\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)}\frac{d\log(\bar{\theta}'_i)}{dp_h} = \beta(\tau_h - \tau_l)\delta\underbrace{\big(p_h\tau_h + (1-p_h)\tau_l\big)}_{=\bar{\tau}}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}. \tag{73}$$

Overall, with $\frac{dp_h}{dC^i_\theta} = \pm\Psi\gamma$, with $\pm = +$ for $x = h$ and $\pm = -$ for $x = l$, one gets (29):

$$\frac{dW(\cdot)}{dC^i_\theta} = \pm\Psi\gamma\beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{74}$$

Lemma 1. Given policy platforms (τ_l, τ_h), an equilibrium of the activism subgame is a set of income group contributions $\{C^l_\theta, C^h_\theta\}$, aggregate contributions (C^l, C^h) defined as $C^i = \int_\theta C^i_\theta d\theta$, probability of high taxes p_h and expected tax rate $\bar{\tau} = E(\tau)$.

To establish existence, start from (28) characterizing optimal contribution of income groups. Sum this expression over income group and get the aggregate effects of activism on political preferences (31).

Using this expression with (25), get (32) which uniquely defines the probability p_h, since the right-hand side of this expression is decreasing in $\bar{\tau}$, hence in p_h. From this probability, one can recover aggregate and individual activism contributions. The probability p_h is unique, because the sensitivity of the education rate to the expected redistribution rate $\bar{\tau}\frac{\epsilon'(\cdot)}{\epsilon(\cdot)} = -\frac{\bar{\tau}}{1-\bar{\tau}}\frac{1}{1+\beta\delta(1-\bar{\tau})} < 0$ depends negatively on $p_h(\tau_l, \tau_h)$, since $\bar{\tau} = p_h\tau_h + (1-p_h)\tau_l$.

The argument for uniqueness goes as follow. Given (τ_h, τ_l), the outcome of the contribution game yield a unique p_h (25). Is there another pair (C^h, C^l) and underlying group contributions that yield the same outcome? No, because the marginal return to contribution is a function of $C^h - C^l$, but the marginal cost depends on individual contribution only, see (28).

Endogenous choice of platforms. To derive (34), note $F(\tau_h, \tau_l, \bar{\tau})$ the right hand side of (32) and let $\mathcal{E}(\bar{\tau}) = \bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ be the elasticity of education rate to the expected tax rate. Totally differentiating (32) w.r.t. p_h and τ_h:

$$\Big[1 - \frac{\partial F(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{dp_h}\Big]dp_h = \Big[\frac{\partial F(\cdot)}{\partial \tau_h} + \frac{\partial F(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{d\tau_h}\Big]d\tau_h. \tag{75}$$

Term by term:

$$\frac{\partial F(\cdot)}{\partial \tau_h} = \Psi\beta\big[(1+\Psi\gamma^2)\sigma'^2(1-\tau_h) + \Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big], \tag{76}$$

$$\frac{\partial F(\cdot)}{\partial \bar{\tau}} = \Psi\beta(\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}. \tag{77}$$

$$\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} = -\frac{1+\beta\delta-\beta\delta\tau^2}{(1-\tau)^2\big[1+\beta\delta(1-\tau)\big]^2} \tag{78}$$

$$\tag{79}$$

Rearranging terms, get (34).

Proposition 3. If $\gamma = 0$ or $\delta = 0$, then the unique solution to (35) is $\tau = 1$. Otherwise, using (54), rewrite (35) as:

$$\Big(1 + \frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1-\tau) = \frac{\tau}{1-\tau}\frac{\delta}{1+\beta\delta(1-\tau)}. \tag{80}$$

The left hand side is decreasing in τ, while the right hand side is increasing in τ, with limit when $\tau = 1$ is ∞, which gives that the unique solution satisfies $0 < \tau^p < 1$. An increase in γ decreases the left hand side, which yield $\frac{d\tau^p}{d\gamma} < 0$. In the limit $\gamma = \infty$, (35) coincides with (14), hence $\tau^p = \tau^*$.

C.3 Citizen-office-seeking candidate

The program of the candidate from party H

$$\max_{\tau_h} \mu p_h + (1-\mu)\Big[p_h\frac{(\tau_h - \tau_h^*)^2}{2} + (1-p_h)\frac{(\tau_l - \tau_h^*)^2}{2}\Big] \tag{81}$$

$$\mu\frac{dp_h}{d\tau_h}+(1-\mu)\Big[\frac{dp_h}{d\tau_h}\frac{(\tau_h-\tau_h^*)^2}{2}+p_h(\tau_h-\tau_h^*)-\frac{dp_h}{d\tau_h}\frac{(\tau_l-\tau_h^*)^2}{2}\Big]=0 \tag{82}$$

Similarly, the first order condition characterizing $\tau_l(\tau_h)$:

$$-\mu\frac{dp_h}{d\tau_l}+(1-\mu)\Big[\frac{dp_h}{d\tau_l}\frac{(\tau_h-\tau_l^*)^2}{2}+(1-p_h)(\tau_l-\tau_l^*)-\frac{dp_h}{d\tau_l}\frac{(\tau_l-\tau_l^*)^2}{2}\Big]=0 \tag{83}$$

To derive these expressions, note $p_h=F(\tau_l,\tau_h,\bar{\tau})$ and get:

$$\Big[1-\frac{\partial F(\cdot)}{\partial\bar{\tau}}\frac{d\bar{\tau}}{dp_h}\Big]dp_h=\Big[\frac{\partial F(\cdot)}{\partial\tau_i}+\frac{\partial F(\cdot)}{\partial\bar{\tau}}\frac{d\bar{\tau}}{d\tau_i}\Big]d\tau_i \tag{84}$$

Then derive

$$\frac{\partial F(\cdot)}{\partial\bar{\tau}}=\Psi\beta(\tau_h-\tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} \qquad \frac{d\bar{\tau}}{dp_h}=\tau_h-\tau_l \tag{85}$$

$$\frac{d\bar{\tau}}{d\tau_h}=p_h \qquad \frac{d\bar{\tau}}{d\tau_l}=(1-p_h) \tag{86}$$

and

$$\frac{\partial F(\cdot)}{\partial\tau_h}=\Psi\beta\big[(1+\Psi\gamma^2)\sigma'^2(1-\tau_h)+\Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big] \tag{87}$$

$$\frac{\partial F(\cdot)}{\partial\tau_l}=-\Psi\beta\big[(1+\Psi\gamma^2)\sigma'^2(1-\tau_l)+\Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big] \tag{88}$$

Finally

$$\mathcal{E}(\tau)=-\frac{\tau}{1-\tau}\frac{1}{1+\beta\delta(1-\tau)} \qquad \frac{d\mathcal{E}(\tau)}{d\tau}=-\frac{1+\beta\delta(1-\tau^2)}{(1-\tau)^2\big[1+\beta\delta(1-\tau)\big]^2} \tag{89}$$